Pitch detection by ACF & AMDF

2017 signal & system

Auto-correlation Function

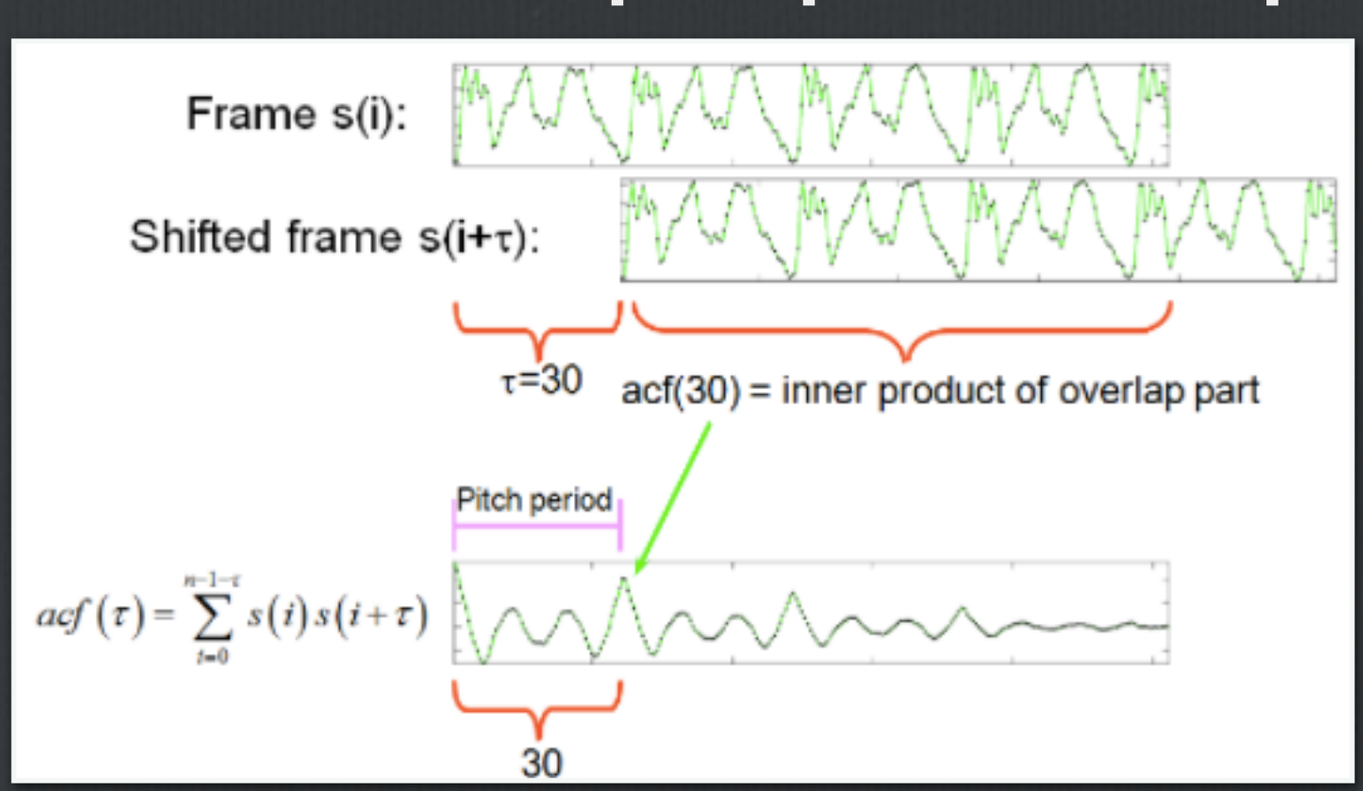
- ACF is a time-domain method which estimates the similarity between a frame s(i),i=0,…,n−1s(i),i=0,…,n−1, and its delayed version via the auto-correlation function:

$$acf(\tau) = \sum_{i=0}^{n-1-\tau} s(i)s(i+\tau)$$

- τ is the time lag in terms of sample points.

ACF

- The value of τ that maximizes acf(τ) over a specified range is selected as the pitch period in sample points.

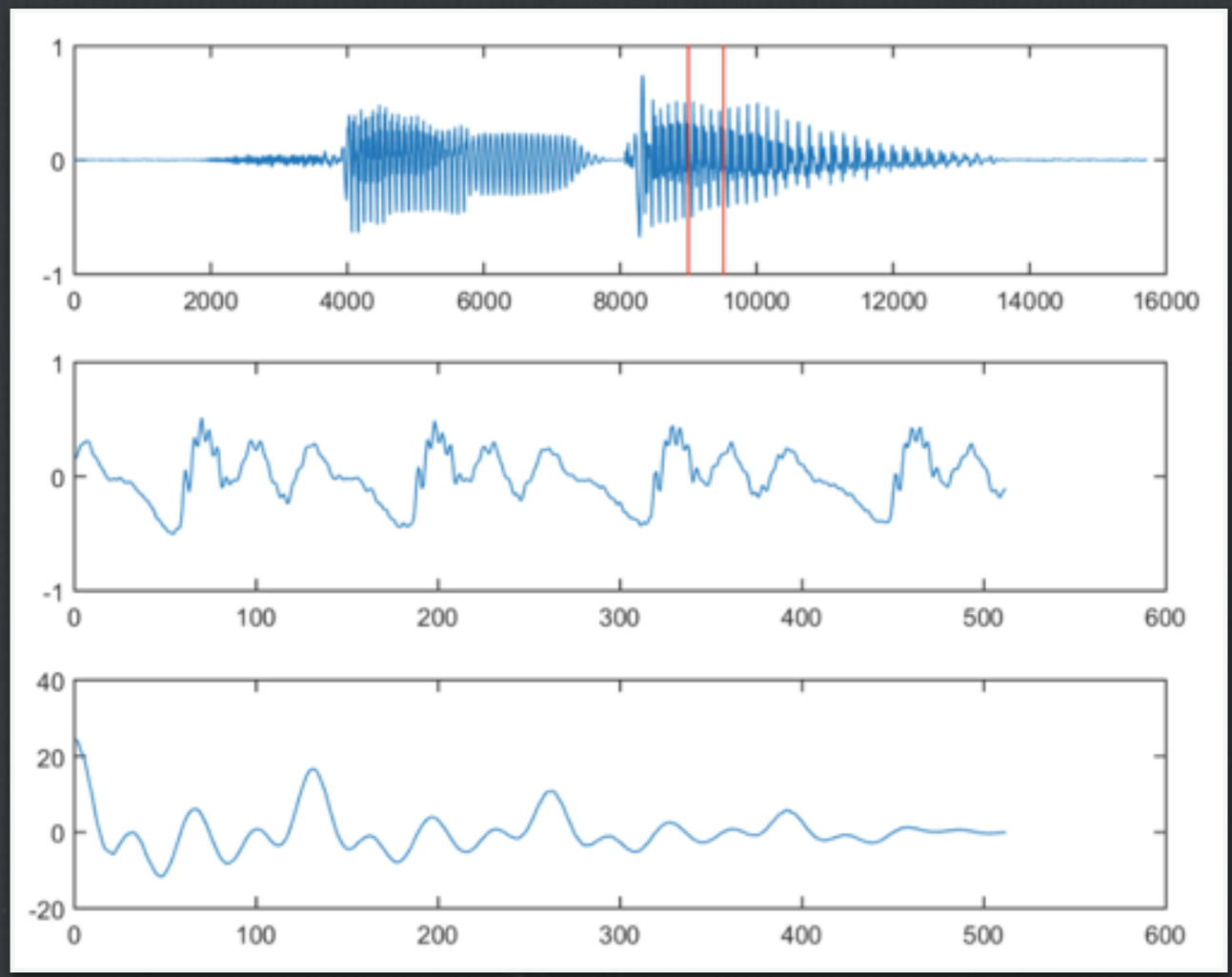


ACF

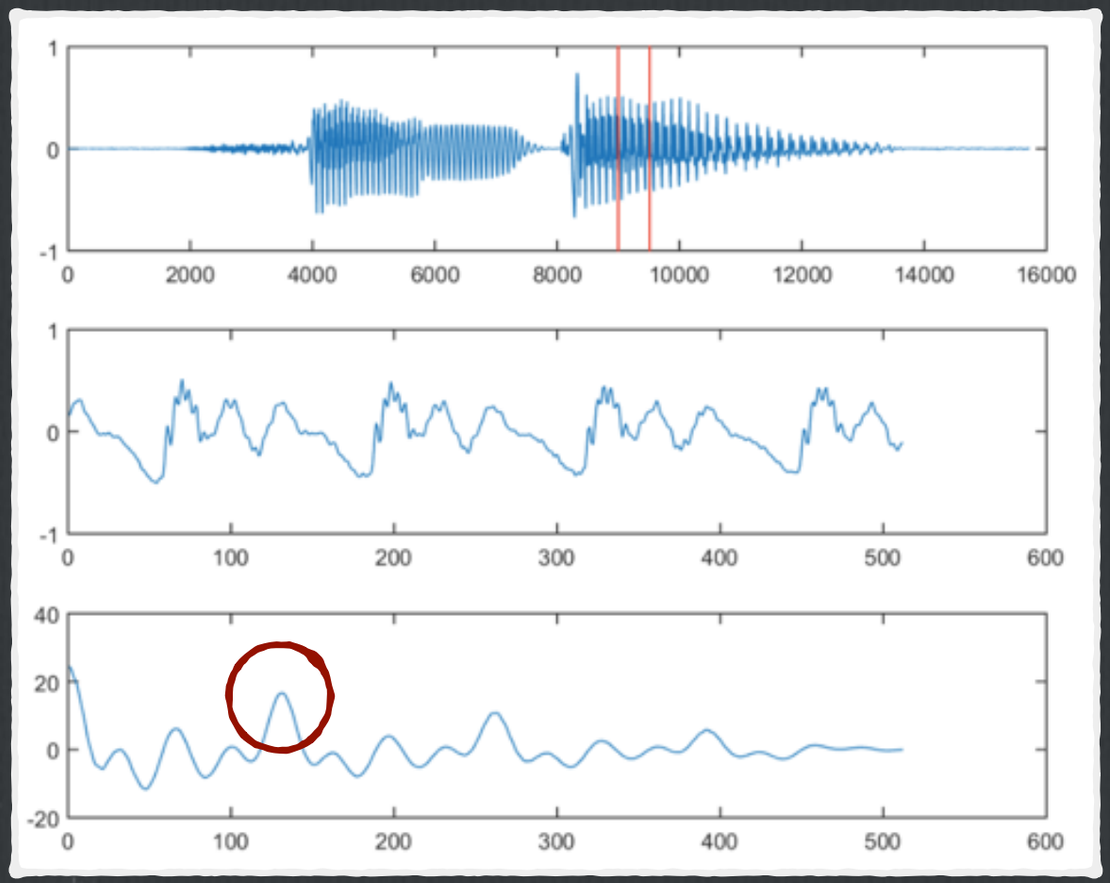
- In other words, we shift the delayed version n times and compute the inner product of the overlapped parts to obtain n values of ACF.

ACF - example

- The maximum of ACF occurs at the first point, which is obviously not what we want.



ACF - example



- Set the values around the first maximum to be $-\infty$, we can identify the second maximum located at index 131 and the corresponding pitch is fs/(131-1) = 16000/130 = 123.08 Hz, or 46.94 semitones.

Average Magnitude Difference Function(AMDF)

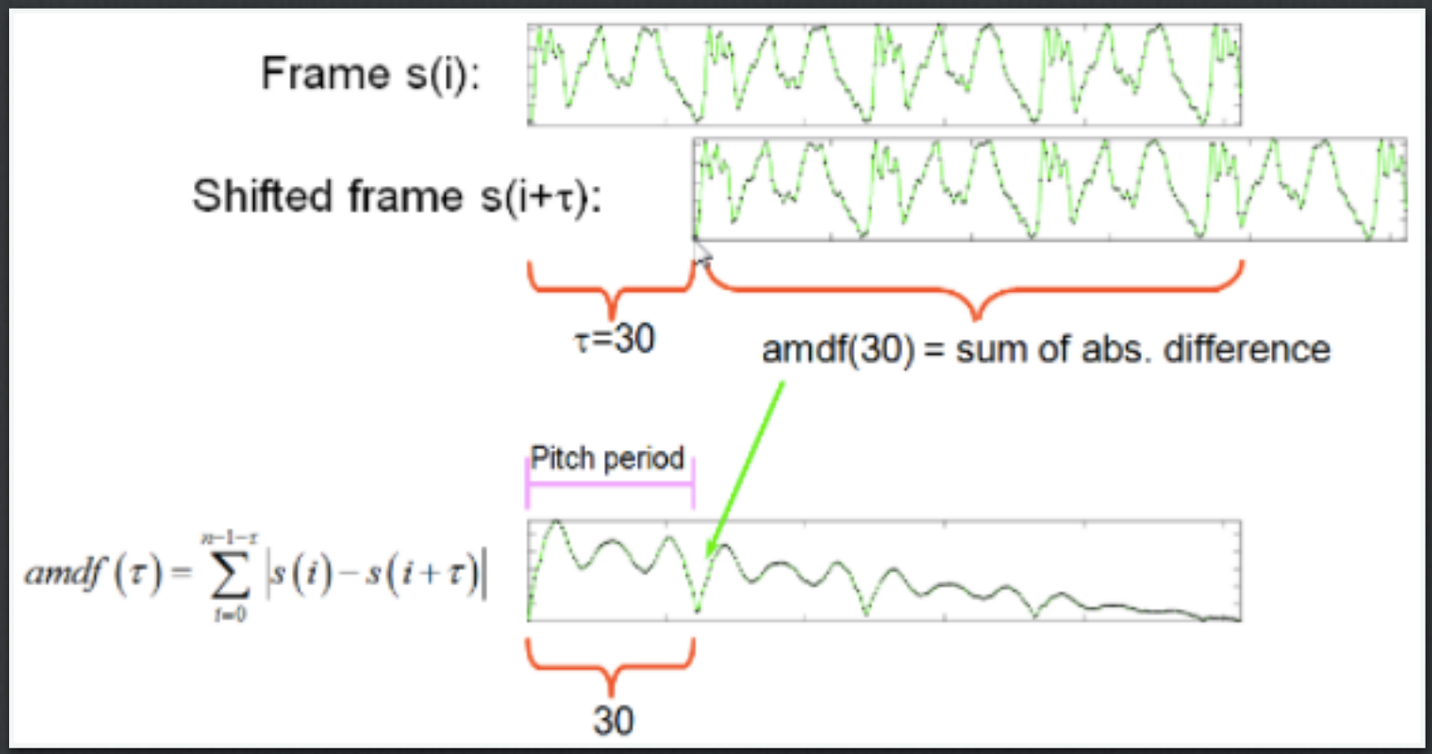
- The concept of AMDF is very close to ACF except that it estimates the distance instead of similarity between a frame s(i), i = 0 ~ n-1, and its delayed version via the following formula:

$$amdf(\tau) = \sum_{i=0}^{n-1-\tau} |s(i) - s(i+\tau)|$$

- τ is the time lag in terms of sample points.

AMDF

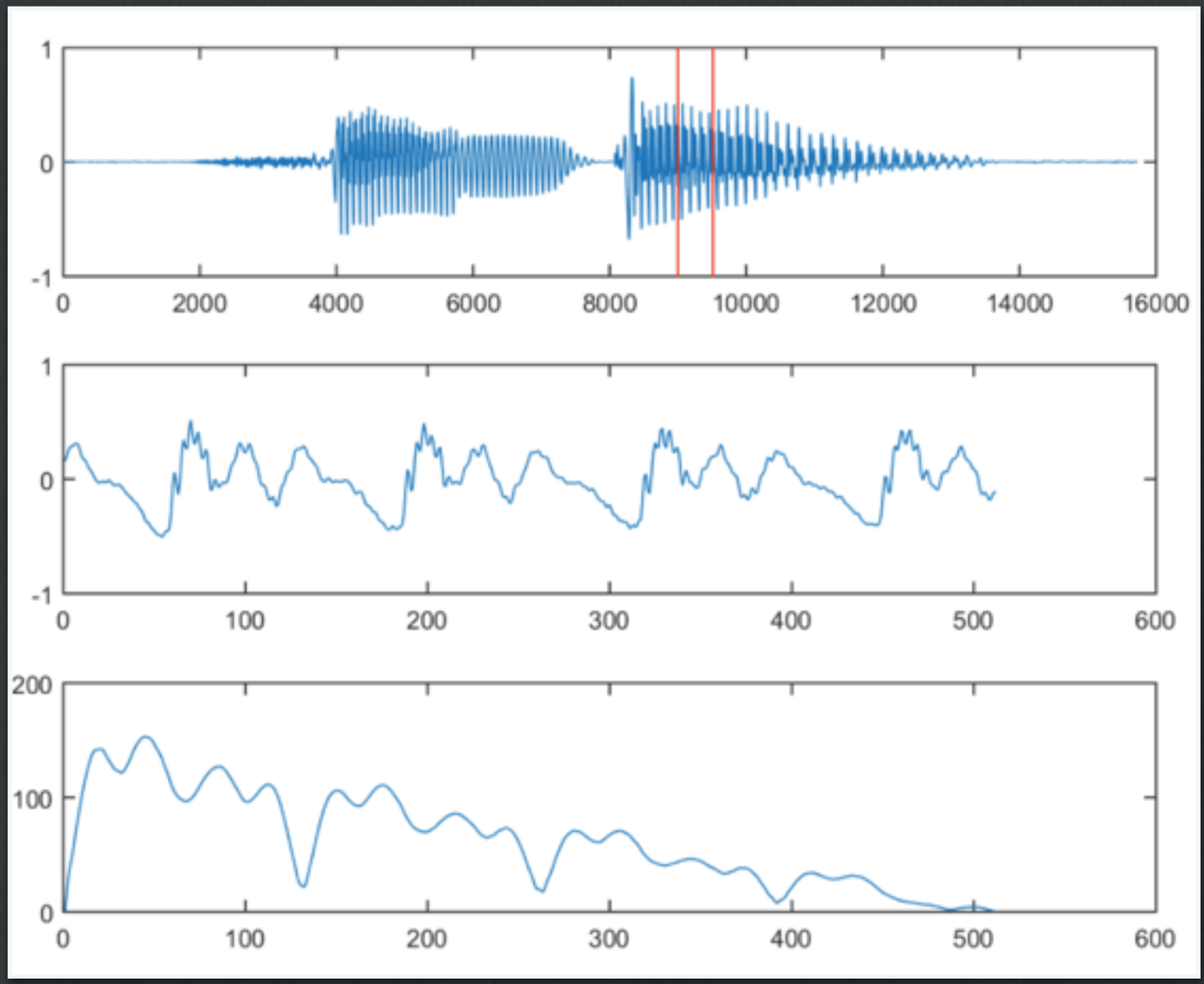
- The value of τ that minimizes amdf(τ) over a specified range is selected as the pitch period in sample points.



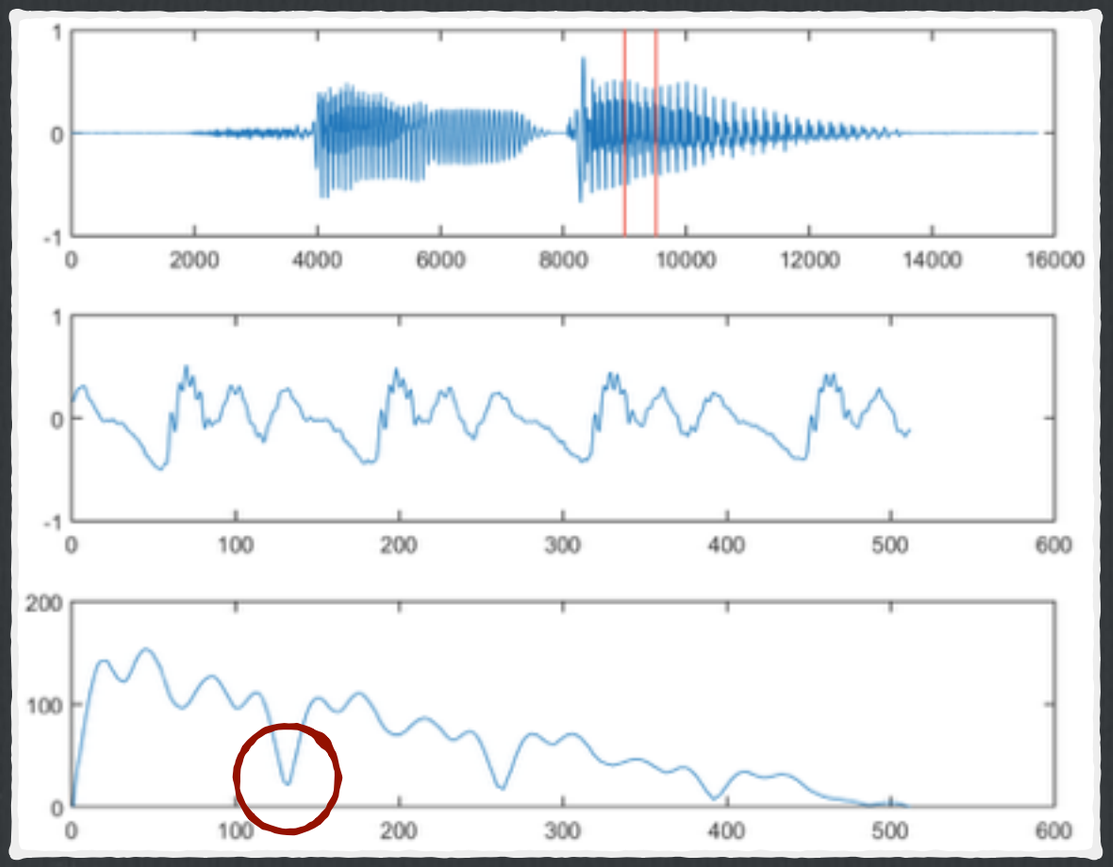
AMDF

- In other words, we shift the delayed version n times and compute the absolute sum of the difference in the overlapped parts to obtain n values of AMDF.

AMDF - example



AMDF - example



- local minimum located at index=132. The corresponding pitch is equal to fs/(132-1) = 16000/131 = 122.14 Hz, or 46.81 semitones. This result is close but not exactly equal to the one obtained via ACF.